

142

MEDICAL RESEARCH LABORATORY



U. S. Naval Submarine Base
New London

A STUDY OF THE INTERRELATIONSHIPS OF PSYCHOLOGICAL
AND PHYSIOLOGICAL MEASURES ON SUBMARINE ENLISTED
CANDIDATES: I. HISTORY, EXPERIMENTAL DESIGN, AND
STATISTICAL TREATMENT OF DATA

Report No. 1 of Bureau of Medicine and Surgery Research
Project No. NM-003-017

Prepared by

Ellsworth B. Cook, Lt. Comdr. (MSC) USN
U. S. Naval Medical Research Laboratory

and

Dr. Robert J. Wherry, Professor of Psychology,
Ohio State University, and Consulting Statistician
to the U. S. Naval Medical Research Laboratory

9 March 1949

APPROVED: T. L. WILLMON, Captain (MC) USN, O-in-C

APPROVED FOR
RELEASE - CONFIDENTIAL

MRL RPT 142

A STUDY OF THE INTERRELATIONSHIPS OF PSYCHO-
LOGICAL AND PHYSIOLOGICAL MEASURES
ON SUBMARINE ENLISTED CANDIDATES

I. HISTORY, EXPERIMENTAL DESIGN, AND
STATISTICAL TREATMENT OF DATA

Report No. 1 of Bureau of Medicine and Surgery Research
Project No. NM-003-017

Report prepared by

Ellsworth B. Cook, Lt. Comdr. (MSC) USN
U. S. Naval Medical Research Laboratory
U. S. Naval Submarine Base, New London, Conn.

and

Dr. Robert J. Wherry, Professor of Psychology,
Ohio State University, Columbus, Ohio, and
Consulting Statistician to Medical Research Laboratory

9 March 1949

Opinions or conclusions contained in this report are those of the authors. They are not to be construed as necessarily reflecting the views or the endorsement of the Navy Department. Reference may be made to this report in the same way as to published articles noting authors, title, source, date, project number and report number.

TABLE OF CONTENTS

	<u>Page No.</u>
ABSTRACT	i
INTRODUCTION:	
Purpose of Paper	1
Statement of the Problem	1
Background of the Problem	1
Present Selection Criteria	3
Possible Additional Indices:	
Stress Tolerance	4
Masculinity	5
Other Factors	5
EXPERIMENTAL PROCEDURE:	
Population	6
Schedule	7
STATISTICAL ANALYSIS OF DATA:	
Statistic Selected	8
Sub-Sets of Data Chosen	9
APPLICATION OF RESULTS	12
- - - - -	
APPENDICES:	
A Experimental Schedule	14
B Rationale for Statistic Used	17
B ₁ - B ₅ Technical Appendices Developing Steps Between Equations of Appendix B	33
C Glossary of Algebraic Symbols Used	39
D Activities Participating in Project	41
E Acknowledgments	44

A B S T R A C T

This report introduces a series of papers to be presented on the interrelationships of some 330 physiological and psychological measures obtained on a population of 120 submarine enlisted candidates.

The purpose and background of the main problem (that of making the submarine selection program more objective) is presented, and current screening measures are discussed together with additional indices included for purposes of comparison.

For convenience in handling the mass of data, the 330 measurements were subdivided into logical subject matter fields and factor analyzed. Individual papers are in preparation on the main areas covered--physical fitness tests, urinary 17-ketosteroid and androgen output and stress tolerance, psychiatric interview, Rorschach, physical characteristics, anthropometric and somatotyping data, blood data and psychological tests. Significant factors which emerged from the area studies will be combined in a final relational study.

This paper is designed to facilitate reading of later reports in the series, and to this purpose, a detailed discussion of the nature of the correlation coefficient and the technique of factor analysis is included in the appendix material.

It is hoped that a small battery of relatively independent tests will be established from the total study, for validation on a subsequent group of subjects prior to suggesting revisions of the submarine selection program. Inasmuch as the variables studied are not peculiar to the submarine service alone, findings should be of interest to the general selection problems of other military organizations and to industry as well.

Area studies have suggested worthwhile leads for further investigations in the individual fields covered. It is anticipated that valuable indications of area interrelationships will be forthcoming from the final relational study.

INTRODUCTION

Purpose of Paper:

This report, the first of a number of papers by this activity on the results of a complex statistical study of the interrelationships of psychological and physiological measures obtained on a naval enlisted population, presents a summation of the investigations included. Herein are discussed the problems which prompted the undertaking and the analytical methods employed to isolate the factors present in the interrelationships studied. Such an understanding is fundamental to a comprehension of later reports in which each area will be treated separately. Details given on statistical procedures used are designed to facilitate interpretation of the data presented in the area reports and will not be repeated.

Statement of the Problem:

The main problem in this investigation is that of improvement of the submarine selection program. To this end attempts were made to evaluate the adequacy of current methods, to study the feasibility of adding or subtracting other measures, and to determine whether significant correlations can be found between tests regularly administered and the techniques added for purposes of this study.

Background of the Problem:

To a degree exceeded nowhere else in the armed forces, the accuracy of selection and placement of submarine personnel is of extreme importance to the welfare of the service, the success of the mission, and the safety of the crew. Despite extensive preliminary screening at other installations prior to their assignment to the Submarine Base, New London, submarine enlisted candidates may be eliminated at three later stages: (1) Medical Research Laboratory processing, (2) Submarine School 8 weeks' basic training

and (3) trial period of six months aboard submarines.

Rejections during stage 1 are for inability to meet physical, psychological and intelligence requirements. Elimination of men as psychologically unsuited is based largely on the results of the psychiatric interview, hence the danger exists of considerable subjectiveness in the qualification or disqualification of candidates.

Rejections during stage 2 arise from scholastic inaptitude, loss of interest in submarine duty, disciplinary reasons, or because, in the opinion of instructors, a man is temperamentally maladjusted for submarines.

Rejections during stage 3 are mainly from lack of incentive to gain facility and knowledge required for final qualification, or for inability to get along with fellow crew members. Final decision as to satisfactory work performance is the responsibility of the commanding officer of each submarine. Cases involving temperamental difficulties are referred to the Squadron Medical Officer for advice.

The fallacies of employing subjective techniques to evaluate the intermediate group which lies between the extremes of any population result inevitably in the admission of some inapt individuals and in the disqualification of others who, given the opportunity, might have made successful submariners. This is true especially during mobilization when sufficient numbers of highly trained interviewers with the requisite submarine experience are not available for the present protracted selection system.

The continuous technical developments in underwater craft create new adjustment problems for personnel and place further burdens on the selection program. For instance, while the use of "snorkel" makes for more effective submarine operation, the prolonged submergence thus permitted imposes additional confinement and stress on crew members. Advances in submarine design must be matched by increasingly accurate objective procedures in the selection of submarine crews.

Present Selection Criteria:

Currently the screening of submarine candidates by the Medical Research Laboratory includes the following measures:

<u>Medical Tests</u>	<u>Sensory Tests</u>
Physical examination	<u>Vision:</u>
Kahn	Night Vision
Urine Analysis	Color Vision
Chest X-ray	*General Vision
Dental examination	
	<u>Audition:</u>
<u>Basic Battery Written Tests</u>	Whispered Voice
General Classification Test,	**Audiogram
Arithmetic, Mechanical Aptitude,	**Pitch and Loudness
Clerical, Mechanical Knowledge,	Discrimination
Personal Inventory	<u>Escape Tank Training</u>
<u>Psychiatric Interview</u>	Two escapes, employing
Individual discussions with	an underwater breathing
submarine medical officer.	device, from the 12', 18'
	and 50' locks of the sub-
	marine escape training
	tank,

Possible Additional Indices for Personnel Selection:

Recent investigations have suggested new avenues for obtaining quantitative estimates of additional factors for possible inclusion in a selection battery of objective tests.

-
- * Visual acuity measurements with Snellen Letter chart at 20' distance; visual acuity and phoria measurements with the Ortho-Rater at the equivalent of 26' and 13" distances,
 - ** Candidates must meet the auditory requirements established for the standing of sound (Sonar) watches.

Stress Tolerance

Tolerance to the physical and emotional stress of long and hazardous war patrols, and to monotony and confinement would appear to be paramount in the make-up of a successful submariner. An adequate measure of such stress tolerance should prove highly predictive of successful adjustment to submarine life.

1. Ketosteroid and Lymphocyte Changes Studies conducted by the Worcester Foundation for Experimental Biology* on several restricted populations indicate the existence of a positive correlation between stress and the output of the urinary 17-ketosteroid substances. Hoagland and Pincus found that the stress undergone by aviation test pilots was reflected in an increase of 17-ketosteroid substances in urine. This relationship was confirmed by additional work employing a long and difficult pursuit meter task as an artificial stress situation. Also it was found that with normal subjects, the 17-ketosteroid increase during stress was accompanied by a decrease in the blood lymphocyte component. However, with psychotics, who were exposed to the pursuit meter task, there was little or no change in 17-ketosteroid output during stress but there was a marked increase in the blood lymphocyte component. Thus, body chemistry seems to offer a means for study of physiological reaction to the stress situation imposed.

This study investigated the 17-ketosteroid and lymphocyte changes induced by (a) taking difficult written tests (referred to hereafter as the psychological stress situation), and (b) undergoing training at the Submarine Escape Training Tank (referred to hereafter as the tank stress situation).

2. Rorschach The Group Rorschach inkblot test of personality was administered to all subjects. In addition to recording the standard scoring for the test, the laboratory

* Bibliographical references will be given in the individual reports covering the various areas studied, and are not included in this general introductory paper.

psychiatrist compiled a special neurotic score as evidence of emotional stress. The Group Rorschach was employed since it appeared to offer greater feasibility for extensive screening programs.

3. Interview The psychiatric interview ordinarily given submarine candidates was more definitely formulated and delineated for this study. Nineteen items were selected in an attempt to combine an evaluation of the personality traits of the individual and his family background with the specific intent of determining measures both of emotional stability and of masculinity. These were graded by two interviewers separately.

Masculinity

The trait of masculinity, while vague and ill-defined, has attracted great interest in selection. This study attempted to evaluate masculinity by the following means:

1. An estimate from physical characteristics: a repetition of the somatotyping work conducted by the Grant Study, Harvard University, was deemed worthwhile in view of the widely differing opinions concerning the possible application of such work to selection problems.
2. A qualitative estimate from interview: as indicated above, the psychiatric interview was more definitely formulated to quantify "masculinity."
3. An index of masculinity from personality tests: the Minnesota Multiphasic Personality Inventory was employed.
4. A biochemical index: urine was analyzed for androgen content and the results were employed as the biochemical measure of masculinity.

Other Factors

While the subjects were available, it was elected worthwhile to include the following additional tests which were

thought to be of possible significance in selection: (1) three physical fitness tests (the Navy Step-up, the Harvard Step-up, and the Schneider); (2) complete differential leucocyte counts; and (3) anthropometric measurements.

The administration of such a variety of tests on a single population presented a unique opportunity for examining the interrelationships in several related but distinct fields.

EXPERIMENTAL PROCEDURE

Population:

The subjects in this study were 120 submarine school candidates. The age range was from 17 to 26 years. They were selected at random and every effort was made to keep motivation high.*

Strict regimentation of the subjects was maintained during their three-day testing period. A secluded section of the laboratory was set aside as barracks space, and a separate mess was provided. A chief hospitalman was assigned as Master-at-Arms to supervise the groups and maintain the schedule.

- - - - -

* Subjects were selected by separating from the main group of approximately 40 candidates the occupants of six chairs which had previously been secretly marked. In the course of an indoctrination lecture by the officer-in-charge of the laboratory, these men were informed that they had been chosen to serve as subjects in an experiment for a three-day period, were impressed with the necessity of complete cooperation during the testing period, and promised compensatory liberty. All subjects were given the opportunity of withdrawing from the experiment; two did so and volunteer substitutes were found. With very few exceptions the men were highly motivated since all desired submarine duty and felt that failure on any of the tests administered would disqualify them for Submarine School.

Schedule:

Two groups of six subjects were studied each week; the first group was selected on Monday morning and the second on Wednesday morning. The schedule for each group was identical, with the exception that the Group Rorschach was administered to both sections simultaneously on Wednesday afternoon.

One physical fitness test was given to the men on each of the three experimental days; these were randomized and administered under the direct supervision of the project medical officers.

One stress situation was given each morning of the second and third experimental days. For the first thirteen groups the psychological stress* was given initially and the tank stress** secondly. The order was reversed for the remaining seven groups since preliminary analysis of data suggested that possibly the tank training produced physical fatigue which might be reflected in lower physical fitness test scores for that day.

* The psychological stress situation consisted of taking the Minnesota Multiphasic Personality Inventory and the Officers Classification Test, with one hour allotted for each. This approximated the mean time customarily allowed the former; the personal nature of the questions contained in this test was considered stressful to the subjects. The Officers Classification Test was considered beyond the general ability of most enlisted candidates and hence more stressful than would be the General Classification Test. Moreover, the hour allotted for completion of the OCT was less than half the normal period permitted. Additional tension was provided by periodically announcing the sections of each test which should have been completed by that particular time interval.

** This consisted of the escape training procedures regularly undergone by submarine candidates. (See page 3)

During the course of each stress period, four urine and blood samples were taken: (1) the basal: the first sample collected upon arising*; (2) the pre-stress: taken 30 minutes before the stress period began; (3) the stress: taken 15 minutes after the stress period ended; and (4) the post-stress: taken 2 hours after the stress period ended.

The complete experimental schedule is presented in Appendix A.

STATISTICAL ANALYSIS OF DATA

Inasmuch as this study was designed to investigate the interrelationships among the measurements taken, the statistic best adapted for this purpose was the correlation coefficient, which indicates the degree and direction of relationship between two variables. If every pair of relationships between the 330 measurements were studied in turn, there would be $(330 \times 329)/2$ or 54,285 such coefficients to compute and evaluate. Many such comparisons were of little interest and their computation would entail needless time and effort. Therefore, the 330 variables were broken into natural sub-divisions, enabling a more meaningful presentation of results by indicating the basically different yet reliable types of measurements within each group, and, finally, the inter-area relationships. Several outstanding statisticians were consulted, together with experts in each of the special subject matter fields. The basic measurements were divided into 8 groups, 3 of which again were sub-divided, making a total of 12 sub-sets of data, each containing from 20 to 47 possibly overlapping and possibly independent measurements from a relatively homogeneous subject matter area. The kind and number of variables selected for each

- - - - -

* In the evening of each experimental day, care was taken to ensure that all subjects emptied their bladders completely before retiring to avoid contamination of the next day's basal ketosteroid sample.

area study were:

	<u>No. of Variables</u>
(1) <u>Physical Fitness</u>	
Navy Step-up Test	6
Harvard Step-up Test	8
Schneider Test	17
Hand Dynamometer	2
Misc. (age, body surface area)	2
Total	<u>35</u>
(2) <u>Ketosteroid Data</u>	
Ketosteroid	8
Creatinine	8
Additional Tests	4
Total	<u>20</u>
(3) <u>Psychiatric Interview</u>	
Interviewer A	19
Interviewer B	19
Composite Scores A	3
Composite Scores B	3
Masculinity Estimate - Grant Study	1
Masculinity Component - Physical Exam.	1
Age in Months	1
Total	<u>47</u>
(4) <u>Physical Characteristics</u>	
General Physical Ratings	25
Age in Months	1
Total	<u>26</u>
(5) <u>Anthropometric Data</u>	
Somatotype Ratings	20
Anthropometric Measurements	13
Masculinity Estimate	1
Disproportions and Age	2
Total	<u>36</u>

(6) Rorschach Data No. of Variables

Rorschach (A)*	
General Items	11
Selected Items	7
Selected Stress Items	8
Total	26
Rorschach (B)*	
General Items	2
Selected Items	17
Selected Stress Items	1
Total	20

(7) Blood Data

A. Lymphocyte and Total White Count
Comparison for Psychological and
Tank Stress:

Total White Count Comparison - Psych	5
" " " " - Tank	5
Lymphocyte Count Comparison - Psych	5
" " " " - Tank	5
Total	20

B. Lymphocyte and Total White Count:

Total White Count	8
Lymphocyte Count	8
Lymphocyte Count Mean %	8
Total	24

C. Leucocyte Count:

Polymorphonuclear	8
Eosinophiles	8
Basophiles	8
Monocytes	8
Total	32

* (A) and (B) represent two independent scorings of the same records.

(7) <u>Blood Data</u> (Cont'd.)	<u>No. of Variables</u>
--	--------------------------------

**D. Leucocyte Comparisons for
Psychological and Tank Stress:**

Polymorphonuclear Comparisons - Psych	3
" " - Tank	3
Eosinophiles Comparisons - Psych	3
" " - Tank	3
Basophiles Comparisons - Psych	3
" " - Tank	3
Monocytes Comparisons - Psych	3
" " - Tank	3
Total	24

(8) Psychological Data

Minnesota Multiphasic Personality Inventory	11
G.C.T. Items	5
Tank	1
Two-Hand Coordination	1
Personal Inventory Test	2
Total	20

Correlation coefficients were obtained for all of the measurements in each area. These provided the basic measures for factor analysis, a statistical technique which explains all the relationships in terms of a relatively small number of factors.

Results of a factorial study give a factor loading for each original measurement on each of the factors isolated. The percent of influence explained by each factor in any particular case is obtained by squaring the factor loading.

Factor loadings never summate exactly to unity due to small chance errors* in the original measurements them-

* Under "chance errors" are subsumed all irrelevant influences both psychological (fluctuations in interest and attention, shifts in emotional attitude, differential effects of memory and practice) and environmental (distractions, noises, interruptions, and so forth.)

selves. The remainders are presented as residual tables. Inspection of these residuals enables the reader to evaluate the thoroughness and effectiveness of the factor analysis and the purity or spuriousness of the original correlations.

The reports in this series, then, will present the factor loadings, and both the original table of intercorrelations and the table of residuals arising when one attempts to predict the correlations by the loadings.

The labeling of a factor is a matter of interpretive judgment rather than a problem in statistics. The designation is assigned on the basis of some property or trait judged to be present in all and only in those tests having loadings on the given factor or reference axis. Insofar as the factor loadings reproduce the correlation coefficients, their mathematical accuracy and certainly their existence cannot be denied, but the name or label assigned is always open to question. While the authors have used their best judgment in identifying the factors, they invite the reader to consider and suggest alternate names.

The particular method of factor analysis employed in this study was the Thurstone centroid method as modified by statisticians of the Personnel Research Section, Adjutant General's Office, Department of the Army. The success of this technique and the soundness of the basic data are both attested by the uniformly low and balanced tables of residuals which are reported in the area papers.

Since reports in this series will be distributed to individuals of widely different fields of specialization, a detailed rationale for the statistical process employed is given in Appendix B, together with a simple illustration of factor analysis procedure.

APPLICATION OF RESULTS

It is hoped that, on the basis of this fundamental study, it will be possible to establish a simple selection battery

consisting of the primary physical, physiological and psychological factors which emerge as significant, and validate them on a second large independent sample of applicants. The results of such a second study, simpler in design than the one reported herein, may then be employed as the means of revising the current selection program.

Inasmuch as the variables examined are not peculiar to a particular branch of the service, it is anticipated that the findings should be of use in the general problems of selection encountered by other military organizations and by industry.

The diversity of the data makes available for the first time the possibility of interrelating, in a single population, measures from the medical, the anthropometrical, the psychological, the psychiatric, the physiological and the biochemical fields. In addition, the individual area studies will suggest possible leads for further research.

Factor analyses of the individual areas are now complete and separate reports are being prepared. Computational work is in progress on the master matrix which will incorporate significant factors from all areas.

A P P E N D I X A

EXPERIMENTAL SCHEDULE

Times given are proximate. Tests employed regularly in submarine selection are indicated by a single asterisk.

First Day:

0745	Report to classroom
0800-0900	*Visual evaluation for radar watches *Psycho-acoustic evaluation for sonar watches **Call Personnel Office for GCT marks on men
0900-0930	Motivation Lecture
0930-1145	*Interviews *Vision tests *Color Vision test *Audition test *Ear, nose and throat examination
1145-1230	Lunch
1230-1300	Assignment of bunks and lockers
1300-1415	*X-ray of chest *Dental examination
1415-1515	Hand Dynamometer test, anthropometric measurements, and (if necessary) completion of *interviews
1515-1530	Rest prior to physical fitness test

** The General Classification Test was not given during the experimental period, but the marks were obtained from the service records of the subjects and used as an additional variable.

1530-1630 Physical fitness test

1700 Dinner

2200 Urinate and discard

2215 Retire

Second Day:

0600 Reveille; collection of urine and blood specimens

0645-0745 Breakfast

0745-0800 Rest period

0800 Collection of urine and blood specimens

0830-1045 Officers Classification Test
*Basic Battery Written Tests
Minnesota Multiphasic Personality Inventory

1045-1100 Rest period

1100 Collection of urine and blood specimens

1130-1215 Lunch

1230-1245 Collection of urine and blood specimens

1300-1500 *Physical examinations
Somatotyping photographs

1500 Rest period prior to physical fitness test

1500-1700 Physical fitness test

1700 Dinner

2200 Urinate and discard

2215 Retire

Third Day:

0600	Reveille; collection of urine and blood specimens
0645-0745	Breakfast
0745-0800	Rest at training tank
0800	Collection of urine and blood specimens at training tank
0800-1100	*Tank training (15 minutes rest seated immediately after completion of training)
1100	Collection of urine and blood specimens
1115-1230	Lunch
1230-1245	Rest period
1245	Collection of urine and blood specimens
1300-1415	Group Rorschach (both sections)
1415-1430	Rest period prior to physical fitness test
1430-1700	**Physical fitness tests
1700	Dinner - Dismissed

** The group which completed its experimental period on Wednesday afternoon and the group which started its session on Wednesday morning each took the physical fitness test allotted it according to the randomized schedule.

A P P E N D I X B

RATIONALE FOR STATISTIC USED

Since factor analysis starts and ends with correlation coefficients, it is necessary to understand thoroughly just what a correlation coefficient means.

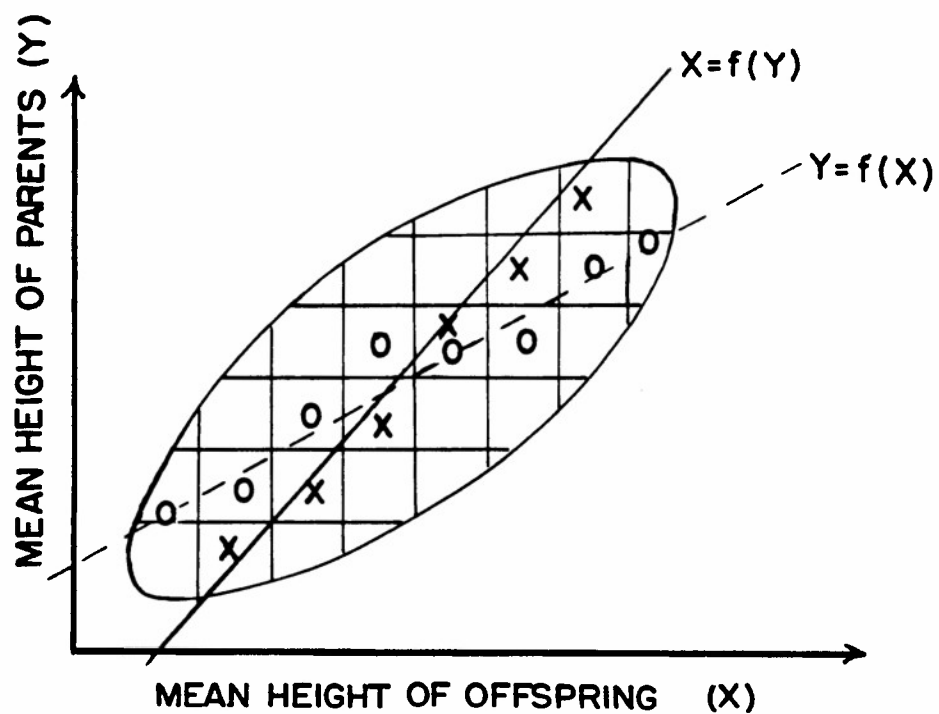
The correlation coefficient was discovered by Sir Francis Galton in connection with his studies on heredity. He was concerned with the relationship existing between the height of parents and their children (when adult). He hit upon the idea of plotting such data in a two way table, now called a double entry table or scatter diagram. He found that dots or tallies representing each pair of measurements tended to fall roughly within the outlines of an ellipse as indicated in Figure I.

The reader will note that the ellipse has been divided into horizontal and vertical strips. The horizontal strips represent the varied heights of the children of groups of parents of relatively homogeneous height, while the vertical strips represent the varied heights of parents of groups of children of relatively homogeneous height. If one wishes to predict the possible mean height of future children for parents of a given height, the best gauge is the mean height of past offspring of parents of the given height. Accordingly, we have identified the mean for each such horizontal strip by placing a cross (X) at that point. The crosses are found to lie (approximately) upon a straight line which has been drawn in and labelled $X = f(Y)$.

Similarly, if one wishes to predict the most probable height of parents for grown children of a given height, one enters circles (O) at the mean point of each vertical strip, and draws the broken line labelled $Y = f(X)$. Noting that these lines predicted that (a) children of exceedingly tall parents will be, on the average, shorter than their parents;

FIGURE I.

DIAGRAMMATIC REPRESENTATION OF A
SCATTER DIAGRAM, SHOWING ELLIPTICAL
FORM AND LINEAR NATURE OF ROW (X)
AND COLUMN (Y) MEANS.



(b) children of exceedingly short parents will be, on the average, taller than their parents; (c) parents of exceedingly tall children will be, on the average, shorter than their children; and (d) parents of exceedingly short children will be, on the average, taller than their children, Galton called this phenomenon regression. The lines $[Y = f(X)]$ and $X = f(Y)$ which demonstrate this property he called regression lines. These regression lines indicate the degree of relationship between the two variables and are the basic fact underlying correlation and factor analysis as well. We shall therefore describe them further.

Since the crosses and circles lie on straight lines, we know that the functional relationship can be expressed in the form

$$\begin{aligned} X &= a + b Y, \text{ and} \\ Y &= a' + b' X \end{aligned} \quad (1)$$

or graphically as shown in Figure II.

Since, however, the a and a' values represent the values of X when Y equals zero and of Y when X equals zero respectively, we can eliminate these terms from the equations if we plot our two variables in terms of scores (called deviation scores) measured from the means of their respective distributions. Thus, letting

$$\begin{aligned} x &= X - M_X, \text{ and} \\ y &= Y - M_Y \end{aligned} \quad (2)$$

we get $x = b y$, and

$$y = b' x \quad (3)*$$

as shown in Figure III, where we have the same lines plotted on a new pair of axes representing the deviation scores.

This indicates that the slopes b and b' of the lines must

* See Appendix B₁ for intermediate steps between equations (1) and (3) above.

FIGURE III.
DEVIATION SCORES REGRESSION LINES.

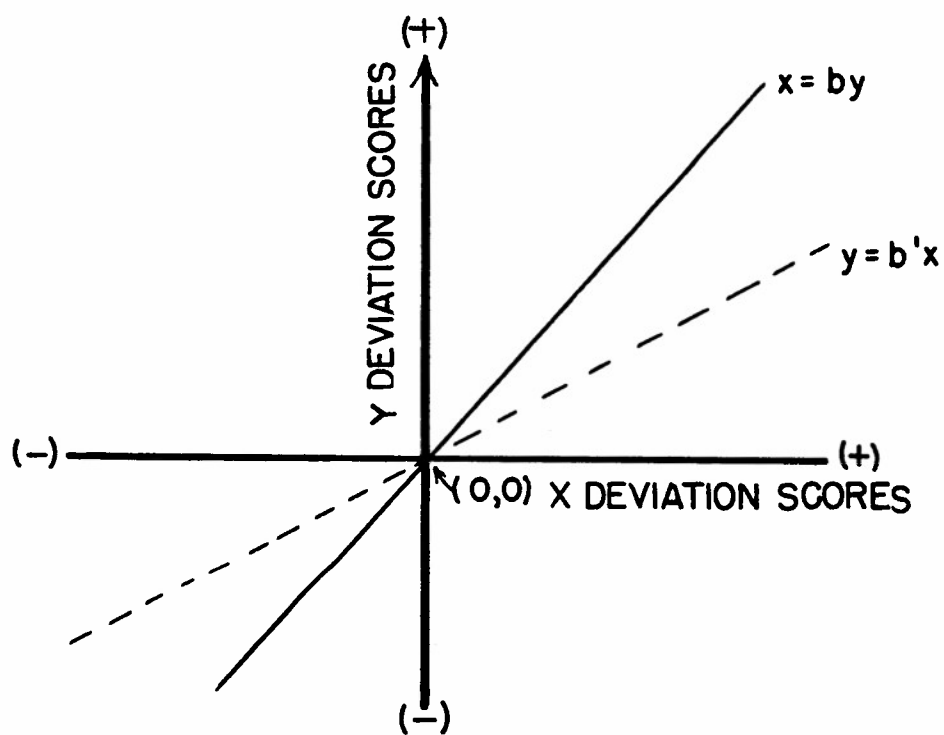
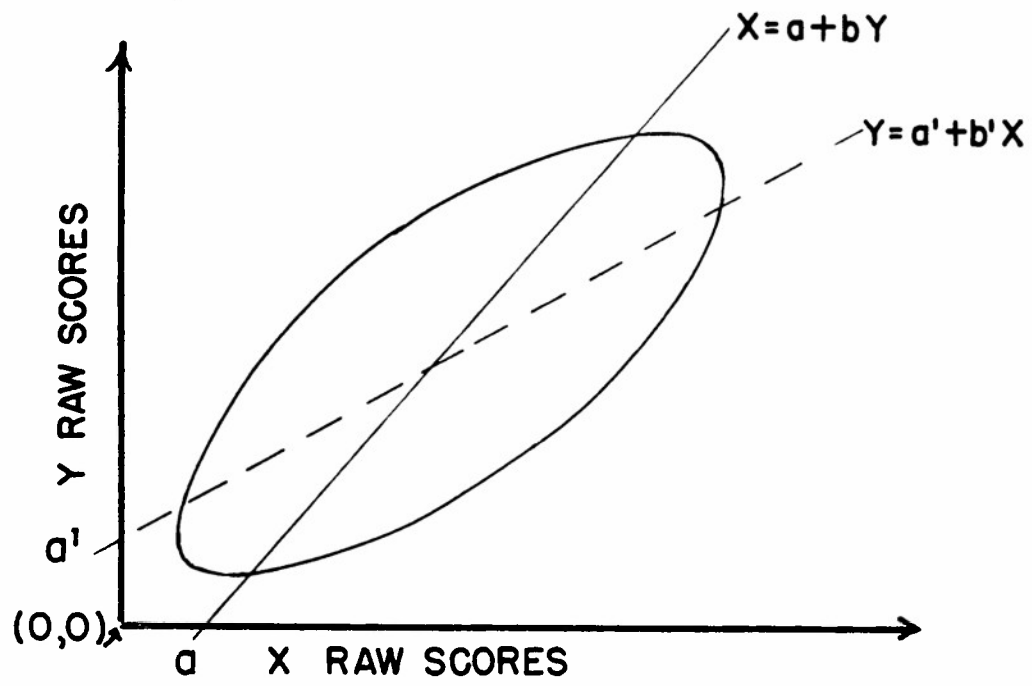


FIGURE II.
RAW SCORE REGRESSION LINES.



be the indicators of relationship. But there are two slopes (b and b^i) and only one relationship. This dilemma was solved by plotting the data and their regression lines on axes representing what the statistician calls standard scores. The cue to this solution was the realization that most variables are plotted in terms of different units of measurement and/or have widely different amounts of relative variability. The standard score is obtained by dividing the deviation score by its own standard deviation, thus eliminating any difference in either unit or variability. Let the symbol z_i indicate the standard score for variable i and we have

$$\begin{aligned} z_X &= x/\sigma_X, \text{ and} \\ z_Y &= y/\sigma_Y \end{aligned} \quad (4)$$

All sets of measurements, in terms of standard scores, have means of zero and standard deviations or variances (squares of the standard deviations) equal to unity. (See Appendix B₂).

The regression lines, in terms of standard scores, take the form

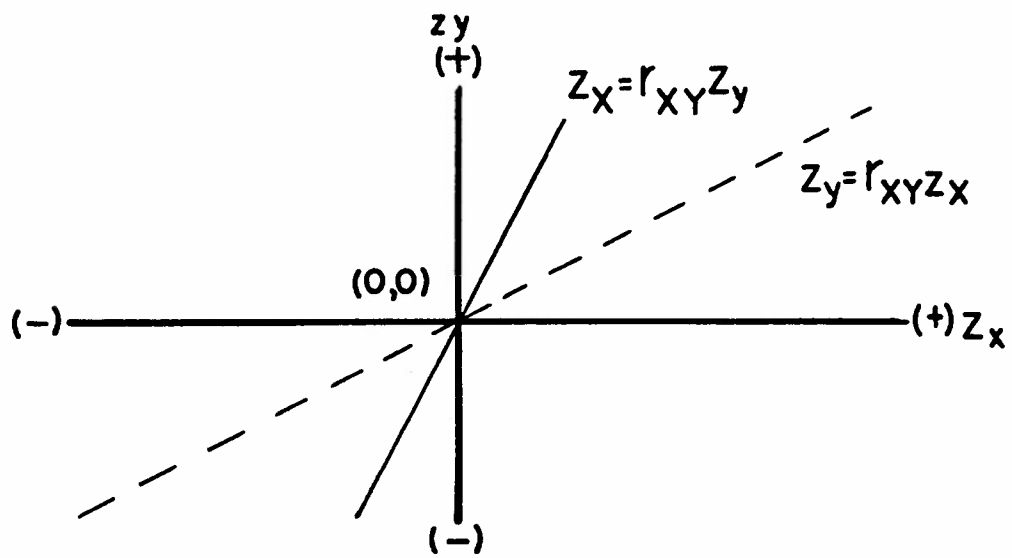
$$\begin{aligned} z_X &= Bz_Y = r_{XY}z_Y, \text{ and} \\ z_Y &= B^1z_X = r_{YX}z_X \end{aligned} \quad (5)$$

(see Appendix B₃) and are illustrated graphically in Figure IV.

That B and B^1 are both equal to the same value, r_{XY} , can be proven by applying the law of least squares and a little calculus of finite differences.* This value of r_{XY} is called the correlation coefficient. Since it appears in both regression equations, it is the single relationship between the two variables which we have been seeking. Karl Pearson, who first discovered the proof of this, called it the product moment correlation coefficient because of the operations performed in obtaining r_{XY} . The correlation coefficients used throughout this series of studies are based upon Pearson's equation where $r_{XY} = \Sigma z_X z_Y / N$.

* See Appendix B₄ for mathematical details.

FIGURE IV.
STANDARD SCORES REGRESSION LINES.



The standard score regression lines have still another function. In addition to being useful in deriving an equation by which the value of r_{XY} can be computed, they serve also to give us a geometric interpretation of the magnitude and direction of the relationship. The coefficient r_{XY} can take all values between + 1.00 and -1.00, and these values are a direct function of the angle between the standard score regression lines. A trigonometric function, the cosine of that angle equals the value of r_{XY} . Consider the graphs in Figure V.

When, as in Figure Va, the ellipse becomes circular, the means for all levels or strips are the same, there is no improvement in prediction, and r_{XY} becomes equal to .00. When, as in Vb and Vc, the ellipse thins down to a straight line and the two regression lines coincide, in this case there is no error of prediction, and r_{XY} is equal to unity in magnitude, positive (+ 1.00) if the relationship is direct, and negative (-1.00) when the relationship is inverse. When the ellipse is neither completely collapsed (to a straight line) nor completely expanded (to a circle), its numerical value lies between ± 1.00 , where the sign (\pm) indicates the direction of magnitude. (See Figures Vd and Ve)

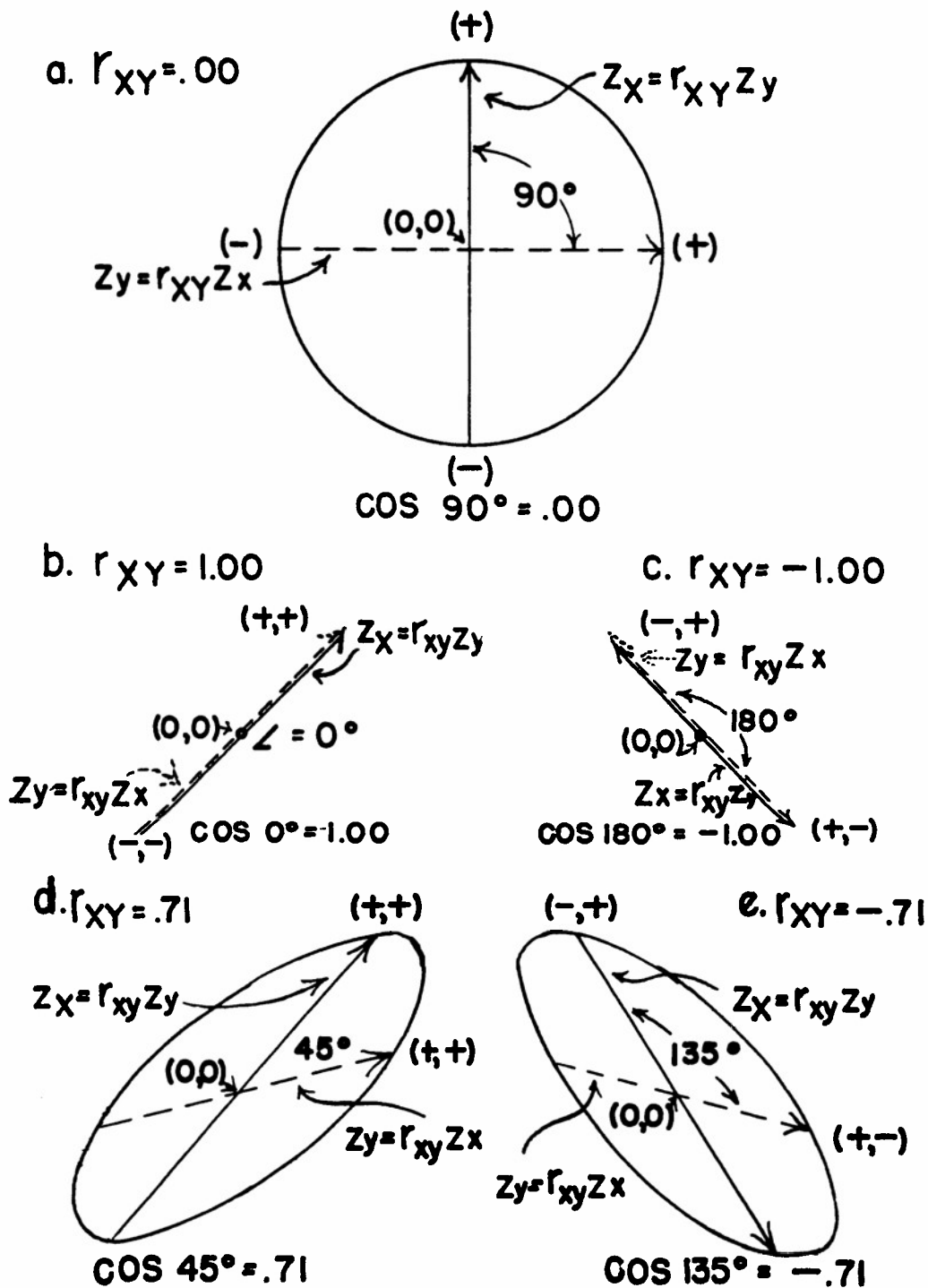
In the preceding paragraph we spoke of the error of prediction and indicated that the magnitude of r (but not its sign) depended upon that error of prediction. This is not strange if one remembers that we used the sum of the squares of the errors of prediction in deriving the formula for r_{XY} . The average of the errors squared is equal to unity minus the square of the correlation coefficient, thus

$$\frac{\Sigma e^2}{N} = 1 - r^2_{XY} \quad (6)$$

(see Appendix B₅) and holds regardless of the direction of prediction. The new value $\Sigma e^2/N$, is called the coefficient of alienation, and tells us what part of the error due to an uninformed guess (prediction of all cases at the mean of the distribution) we still have when we use the other variable (through the regression line) to improve our prediction. Since the value of $1 - r^2_{XY}$ tells us what we have not pre-

FIGURE V.

DIAGRAMS SHOWING RELATIONSHIP
BETWEEN MAGNITUDE OF CORRELA-
TION COEFFICIENT AND THE ANGULAR
SEPARATION OF THE REGRESSION LINES.



dicted, it follows that r_{XY}^2 , called the coefficient of determination, tells us what proportion of the error of prediction has been eliminated. Thus an r_{XY} of .50 will determine .25 of the variance of z_X when predicted by z_Y (or vice versa) and leaves $1 - .25$ or .75 of the variance still unexplained or unpredicted. Later we shall see that factor loadings representing correlations between a given measurement and one of the factors, can be interpreted in this same fashion. That is to say, the square of a factor loading tells us what proportion of the variance of a test is explained or predicted by the factor in question.

We return now to the conception of the correlation coefficient as the cosine of the angle between the axes representing the regression lines. This conception permits us to conceive of any table of intercorrelation among variables as being represented by a series of axes. Each test in this space will be represented by a vector (a line of a given length and direction) of unit length. The unit length 1.00 (remember that $\Sigma z^2/N$ always equals 1.00) indicates the total standard score variance of 1.00 for each test. (See Appendix B₂). The directions of the vectors will be determined by the correlation among the tests. Thus, the angle between the tests 1 and 2 will be the angle whose cosine is equal to r_{12} , and so on for all the remaining tests. The origin of these vectors (the 0,0,0,0,...,0 point for all tests) is chosen arbitrarily as any point not represented by any of the tests in the battery. The position of the first test vector is also completely arbitrary, but the remaining vectors must be put into the system exactly as prescribed by the correlation coefficients with all other tests already represented. The mathematical space in which such vectors are plotted must not be confused with popular physical space which is limited to the three dimensions of height, width and depth. The dimensions here are rather those which determine the variance and covariance (correlation) within and among the tests being pictured. The number of dimensions is not restricted, being as numerous as is required to include all of the test vectors with the required angles between them (the number of dimensions is usually greater than 3).

Factor analysis tells us how many statistically independent (orthogonal or right-angled) axes, called reference axes, or factors, are needed to represent all of the relationships among a set of tests. Obviously, if all measures correlated + 1.00 with each other, all of the vectors would lie along a single reference axis and there would be a single factor underlying all of the tests. Again, if all variables correlated .00 with each other, the vectors for all tests would already be at right angles (orthogonal) to each other, and each test would itself represent a reference vector or factor, and there would be no common reference axis or factor.

If we represent the above extreme cases algebraically we would have, as regression equations, for the case where all r_{ij} values equalled + 1.00,

$$\begin{aligned} z_1 &= 1.00z_A \\ z_2 &= 1.00z_A \\ z_3 &= 1.00z_A \\ &\vdots \\ z_n &= 1.00z_A \end{aligned} \quad (7)$$

where A represents the common reference axis or ability with which the tests are completely identified, and z_A is the standard score of any given individual in reference ability A. In the case where all r_{ij} values equalled .00, we would have

$$\begin{aligned} z_1 &= 1.00z_{S_1} + .00z_{S_2} + .00z_{S_3} + \dots + .00z_{S_n} \\ z_2 &= .00z_{S_1} + 1.00z_{S_2} + .00z_{S_3} + \dots + .00z_{S_n} \\ z_3 &= .00z_{S_1} + .00z_{S_2} + 1.00z_{S_3} + \dots + .00z_{S_n} \\ &\vdots \\ z_n &= .00z_{S_1} + .00z_{S_2} + .00z_{S_3} + \dots + 1.00z_{S_n} \end{aligned} \quad (8)$$

where the reference axes $S_1, S_2, S_3, \dots S_n$ were identical with the original orthogonal vectors representing and specific to the individual test represented by the subscripts 1, 2, 3, ..., n.

While inadvertently we might collect the same information under different names (equation 7 above) for a few variables, or might collect a few measurements completely unrelated (equation 8 above) to each other if the fields of measurement are sufficiently diverse, neither situation is apt to develop in a practical case, especially if the measurements collected be chosen deliberately to represent different aspects of some area related in a logical manner.

Thurstone has stated that the most likely case, to be found under the above circumstances, would yield a set of regression equations of the form

$$\begin{aligned} z_1 &= L_{1A}z_A + L_{1B}z_B + \dots + L_{1R}z_R + L_{1S_1}z_{S_1} + L_{1E_1}z_{E_1} \\ z_2 &= L_{2A}z_A + L_{2B}z_B + \dots + L_{2R}z_R + L_{2S_2}z_{S_2} + L_{2E_2}z_{E_2} \\ &\cdot \\ &\cdot \\ &\cdot \\ z_n &= L_{nA}z_A + L_{nB}z_B + \dots + L_{nR}z_R + L_{nS_n}z_{S_n} + L_{nE_n}z_{E_n} \end{aligned}$$

where -

- (1) the L_{ij} values represent in turn the loadings or correlations of the i th test with the j th factor or reference axis;
- (2) the factors A through R represent orthogonal factors or reference axes upon which two or more of the tests have loadings other than zero;
- (3) the factors $S_1, S_2, S_3, \dots S_n$ represent other orthogonal reference axes upon which only the test of that subscript has a loading other than zero, and which are all orthogonal to the axes A through R as well, but which represent reliable repeatable parts of the specific tests; and

- (4) the factors $E_1, E_2, E_3, \dots, E_n$ represent still other reference axes (orthogonal to each other, the A through R , and the S_1 through S_n axes) upon which only the test of that subscript has a loading and which represent unreliable, non-repeatable or error portions of the individual measurements.

Let us assume specific values of the L_{ij} coefficients for a set of three measurements having two common factors, A and B . Thus we might have

$$\begin{aligned} z_1 &= .6 z_A + .4 z_B + .6 z_{S_1} + \sqrt{.12} z_{E_1} \\ z_2 &= .2 z_A + .8 z_B + .5 z_{S_2} + \sqrt{.07} z_{E_2} \\ z_3 &= .0 z_A + .3 z_B + .9 z_{S_3} + \sqrt{.10} z_{E_3} \end{aligned} \quad (10)$$

Now, from our knowledge of the equation for a correlation coefficient we can write

$$\begin{aligned} r_{12} &= \frac{\sum z_1 z_2}{N} = \frac{\sum (.6z_A + .4z_B + .6z_{S_1} + \sqrt{.12}z_{E_1})(.2z_A + .8z_B + .5z_{S_2} + \sqrt{.07}z_{E_2})}{N} \\ &= \left[.12 \frac{\sum z_A^2}{N} + .32 \frac{\sum z_B^2}{N} \right] + \left[.48 \frac{\sum z_A z_B}{N} + .30 \frac{\sum z_A z_{S_2}}{N} + .6 \sqrt{.07} \frac{\sum z_A z_{E_2}}{N} + .08 \frac{\sum z_A z_B}{N} \right. \\ &\quad + .20 \frac{\sum z_B z_{S_2}}{N} + .4 \sqrt{.07} \frac{\sum z_B z_{E_2}}{N} + .12 \frac{\sum z_A z_{S_1}}{N} + .48 \frac{\sum z_B z_{S_1}}{N} \\ &\quad + .30 \frac{\sum z_{S_1} z_{S_2}}{N} + .6 \sqrt{.07} \frac{\sum z_{S_1} z_{E_2}}{N} + .2 \sqrt{.12} \frac{\sum z_A z_{E_1}}{N} \\ &\quad \left. + .8 \sqrt{.12} \frac{\sum z_B z_{E_1}}{N} + .5 \sqrt{.12} \frac{\sum z_{E_1} z_{S_2}}{N} + \sqrt{.0084} \frac{\sum z_{E_1} z_{E_2}}{N} \right] \end{aligned} \quad (11)$$

But all of the terms in the last bracket represent correlations between orthogonal axes whose correlations are .00 by definition, and hence the whole second bracket reduces to zero. The two terms in the first bracket, however, contain the variances of the identical axes A and B which are equal to unity, and hence $r_{12} = .12 (1.00) + .32 (1.00) + .00 = .44$. Thus we see that the factor loadings on the common factor axes determine the correlation between variables, as well as describing the relation of the test or measurement to those axes. Thus, in general

$$r_{iJ} = L_{iA}L_{JA} + L_{iB}L_{JB} + L_{iC}L_{JC} + \dots + L_{iR}L_{JR} \quad (12)$$

If, on the other hand, we apply the loadings to the problem of the correlation of the test with itself, rather than with another variable, we have

$$\begin{aligned} r_{ii} &= \frac{\sum z_i z_i}{N} = \frac{\sum z_i^2}{N} = 1.00 \\ &= \frac{\sum (L_{iA}^2 A + L_{iB}^2 B + \dots + L_{iR}^2 R + L_{iS_i} z_{S_i} + L_{iE_i} z_{E_i})^2}{N} \\ &= \left[L_{iA}^2 \frac{\sum z_A^2}{N} + L_{iB}^2 \frac{\sum z_B^2}{N} + \dots + L_{iR}^2 \frac{\sum z_R^2}{N} \right] + L_{iS_i}^2 \frac{\sum z_{S_i}^2}{N} + L_{iE_i}^2 \frac{\sum z_{E_i}^2}{N} \\ &\quad + \left[2L_{iA}L_{iB} \frac{\sum z_A z_B}{N} + 2L_{iA}L_{iC} \frac{\sum z_A z_C}{N} + \dots + 2L_{iS_i}L_{iE_i} \frac{\sum z_{S_i} z_{E_i}}{N} \right] \end{aligned} \quad (13)$$

where, as above, the values of $\sum z_i^2/N$ in the first bracket and the next two lone terms all equal 1.00; and the values of $\sum z_i z_j/N$ in the last bracket all equal .00 since all reference axes are defined as being orthogonal. Hence r_{ii} (self-correlation) = $\sum_{J=A}^{J=R} L_{iJ}^2 + L_{iS_i}^2 + L_{iE_i}^2 = 1.00$, which is merely

to say that a set of measures, including errors, are identical with themselves. (This explains, however, how the coefficients with the square root signs in the example above

were estimated).

If we repeated the set of measurements later upon the same set of people, the common factors and the specific factor (S_i) would persist, but the error portion of the new measurements would be unrelated (be orthogonal to the old error axis), thus

$$z_i' = L_{iA}z_A + L_{iB}z_B + \dots + L_{iR}z_R + L_{iS_i}z_{S_i} + L_{iE_i}'z_{E_i} \quad (14)$$

and following the same algebraic development, as above, we would get

$$r_{ii}' = \sum_{J=A}^{J=R} L_{iJ}^2 + L_{iS_i}^2 \approx 1.00 \quad (15)$$

which indicates how reliably a given measure would be repeated and is called the reliability coefficient.

Another interesting way of looking at self-correlation is to separate out the coefficients into two portions, one composed of correlations due to common factors, and the other composed of specific plus error components, thus

$$\begin{aligned} r_{ii} \text{ (self-correlation)} &= \sum_{J=A}^{J=R} L_{iJ}^2 + (L_{iS_i}^2 + L_{iE_i}^2) \\ &= h_i^2 + u_i^2 \end{aligned} \quad (16)$$

where h_i^2 is called the communality coefficient and u_i^2 the uniqueness of the test. It might be noted that h_i^2 corresponds to the coefficient of determination (variation accounted for by common factor space) while u_i^2 corresponds to the coefficient of alienation (variance not accounted for by common factor space).

By way of tying together the algebraic treatment above with the geometric (trigonometric) interpretation of the correlation coefficient given earlier, consider the diagrammatic representation of the relationship between two variables

given in Figure VI. Let us take, for example, an apparently nonsense correlation of .48 between effectiveness of state school programs and sale of chewing gum in the various states. Obviously, any attempt to explain either variable in terms of the other appears ridiculous. However, the correlation does indicate some overlap in factor structure. The Thurstone concept of explanation in terms of common plus specific factors yields a very satisfactory and logical explanation of the empirical relationship. For example (assuming error to be zero, in order to picture the whole situation in 3 dimensions and thus make possible Figure VI), we might have

$$\begin{aligned} z_{\text{School System}} &= L_{(ss)}A z_A + L_{(ss)}S_{ss} z_{S_{ss}} \\ z_{\text{Chewing Gum}} &= L_{(cg)}A z_A + L_{(cg)}S_{cg} z_{S_{cg}} \end{aligned} \quad (17)$$

Our limiting conditions are, therefore, that our solution satisfy

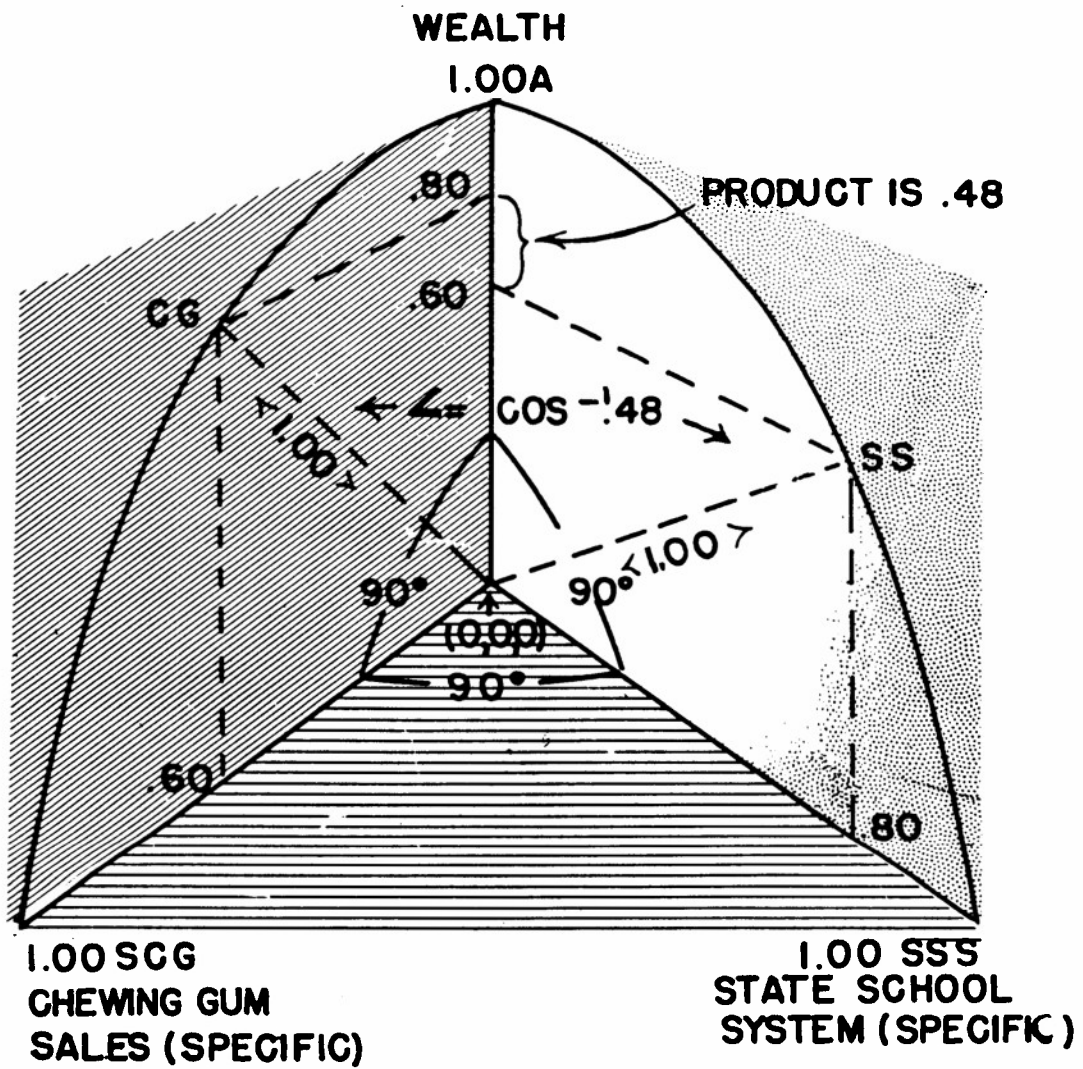
$$\begin{aligned} L_{(ss)}A L_{(cg)}A &= .48 \\ L_{(ss)}^2 A + L_{(ss)}^2 S_{ss} &= 1.00, \text{ and} \\ L_{(cg)}^2 A + L_{(cg)}^2 S_{cg} &= 1.00 \end{aligned} \quad (18)$$

Hence, one acceptable solution would be

$$\begin{aligned} z_{ss} &= .6 z_A + .8 z_{S_{ss}} \\ z_{cg} &= .8 z_A + .6 z_{S_{cg}} \end{aligned} \quad (19)$$

(See Appendix B₅) where the common axis A represents wealth, which determines available expenditures for both school systems and chewing gum and we have assumed the latter (a luxury) to be more affected than the former (a semi-

FIGURE VI.
SPATIAL REPRESENTATION OF A
FACTOR STRUCTURE.



necessity). These values satisfy the necessary requirements and would be diagrammatically represented by Figure VI, shown on the preceding page. This example illustrates another reason for the use of factor analysis in the present data. A single isolated correlation coefficient is always meaningless. Factor analysis is the only tool available for summarizing these individual correlations into meaningful underlying relationships which enable us to understand and interpret such isolated associations.

In actual problems we usually know only the correlations among the variables, and these will yield only the loadings or projections of the tests upon the common reference axes. If in addition we know the reliability of the given tests we can find the specific loadings by taking the square root of the difference between the communality and reliability coefficients ($\sqrt{r_{ii'} - h_{ij}^2}$) and the error loadings by taking the square root of the difference between the reliability coefficient and unity ($\sqrt{1 - r_{ii'}}$).

In the present series of studies the reliabilities are usually unknown and so we will report only the common factor loadings for the variables. We can, however, always obtain a lower limit estimation of the reliability coefficient by computing the communality coefficient which will always be equal or smaller.

A problem encountered in practice is caused by small chance errors in the correlation coefficients themselves. This means that the factor loadings will not exactly reproduce the correlation coefficients upon which they are based. This difference between the sum of the cross products of the common factor loadings for two variables and their actual correlation coefficient is called a residual.

Inspection of the residual tables enables one to evaluate the thoroughness of the statistical treatment. The residuals should be scattered symmetrically about zero (being equally often positive and negative if due to chance) and their range and variability should be as small as possible. Any non-zero grouping or patterning of the residuals would indicate

an incomplete or undefined factor analysis. Scattered but unpatterned large residuals would indicate high chance errors in the original correlation coefficients.

For purposes of illustration, a simple example of this procedure is shown for some tests of vision using an abbreviated problem from an earlier joint Army-Navy study by the authors.

Seven tests of vision were administered to a population of several hundred young men individually in a 20-foot testing alley:

- (1) Bausch and Lomb Checkerboard Test (Adapted)
- (2) New London Letter Test
- (3) Army Air Force Letter Test
- (4) Army Snellen Letter Test
- (5) Dot Variable Size Test
- (6) Line Variable Thickness Test
- (7) Quadrant Contrast Test

The intercorrelations of the seven tests were:

	01	02	03	04	05	06	07
01	-	.75	.77	.78	.78	.74	.48
02		-	.82	.81	.79	.73	.54
03			-	.84	.83	.76	.54
04				-	.83	.82	.61
05					-	.85	.65
06						-	.69
07							-

A factor analysis yielded three factors with the following loadings:

Test	Factor A	Factor B	Factor C	h^2
1	.90	-.01	-.03	.81
2	.84	.11	.21	.76
3	.89	.09	.29	.88
4	.85	.28	.24	.86
5	.87	.31	.08	.86
6	.82	.44	-.01	.87
7	.55	.55	-.01	.61

When the theoretical correlations were computed on the basis of these loadings in accordance with the usual procedure, as follows,

$$r_{12} = (.90)(.84) + (-.01)(.11) + (-.03)(.21) = .75$$

$$r_{13} = (.90)(.89) + (-.01)(.09) + (-.03)(.29) = .79$$

et cetera, the following residual table was obtained:

	01	02	03	04	05	06	07
01	-	.00	-.02	.02	.00	.01	-.01
02		-	.00	.01	.01	-.01	.02
03			-	-.01	.00	-.01	.00
04				-	-.02	.00	-.01
05					-	.00	.00
06						-	.00
07							-

Since the residual table seemed satisfactory, the next step was an attempt to interpret the factors A, B and C.

The interpretations were as follows:

Factor A - This factor was identified as Retinal Resolution due to its very high loadings on all of the tests, especially 1 through 6, since all of the first six tests were designed to measure retinal resolving power. While test 7 was designed primarily as a test of brightness discrimination, its relatively small, although constant, dimension would still require some resolving power.

Factor B - This factor was identified as Brightness Discrimination in part because its highest loading was on test 7 designed to measure that function. The smaller loadings on the dot and line tests would appear to be due to ability of subjects who failed to focus (and thus resolve the dot or line as such) to still give correct responses by noticing a darker area in one corner due to minimal or summational effects as the dot or line moved over a relatively larger portion of the retina. The appearance of a barely significant loading on only the Snellen of the three letter tests is perhaps accounted for by the work done to improve the letters on the Navy and Air Force charts. Since the checkerboard areas and solid gray areas on the checkerboard test (1) had been equated for brightness, this test had to have a zero loading.

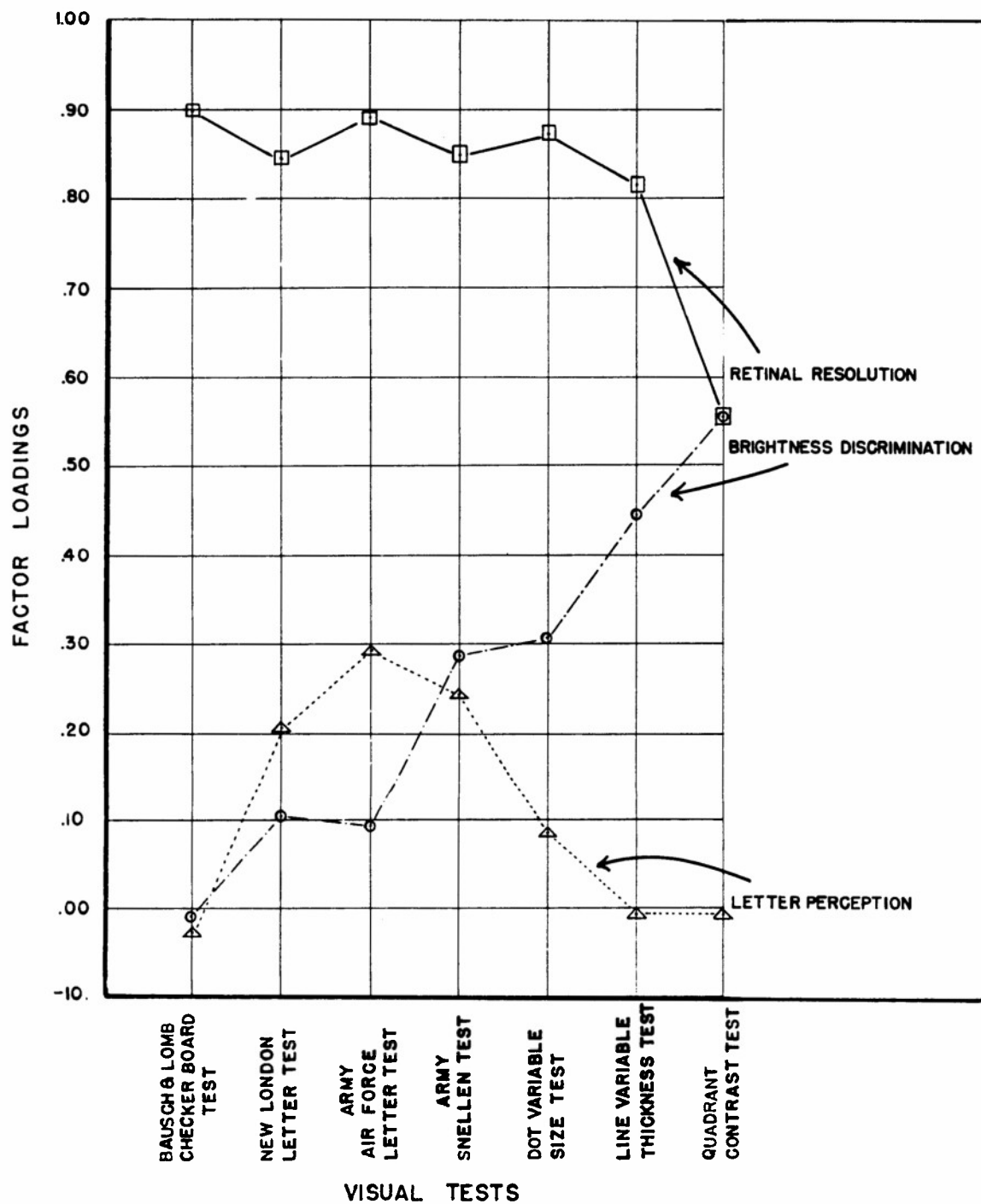
Factor C - This factor, with loadings on all three, and only on the letter tests (2, 3, 4) is called Letter Perception. It appears to be some central perceptual process rather than an optical one per se.

An inspection of the h^2 column shows communalities ranging from .61 to .88, indicating that reliabilities (since they must be even higher) are in general satisfactory. The highness of the communalities indicates also that these three factors account rather adequately for all behavior exhibited on the tests, leaving relatively little to be accounted for by specific or error behavior. For purposes of illustration, the loadings for this study are presented graphically in Figure VII.

FIGURE VII

GRAPHICAL REPRESENTATION OF ROLE PLAYED BY THREE FACTORS (RETINAL RESOLUTION, BRIGHTNESS DISCRIMINATION AND LETTER PERCEPTION) IN SEVEN VISUAL TESTS.

A. FACTOR LOADINGS



The main conclusion to be drawn from the study is that the modified Bausch and Lomb checkerboard test is the sole one which measures resolution and only resolution, and is, therefore, to be preferred over the other measures of far visual acuity investigated.

Each of the sub-sets of data to be presented in the series of reports to follow this introductory paper, will be analyzed in a manner paralleling that of the visual material presented above for purposes of illustration. Familiarity with this section on statistical analysis will enable the prospective reader of these future reports to assimilate their findings with greater ease.

TECHNICAL APPENDIX B₁

As shown on page 18 (Appendix B) the formulas for the lines of regression are:

$$X = a + b Y$$

$$Y = a' + b' X$$

In terms of deviation scores these equations become:

$$x = b y$$

$$y = b' x$$

The proof follows:

$$\text{Let } x = X - M_X \text{ (deviation score)}$$

$$y = Y - M_Y$$

$$\text{Then } X = x + M_X$$

$$Y = y + M_Y$$

and noting that $X = a + b Y$ can be written as $X - a - bY = 0$

summing for all cases this becomes $\Sigma X - Na - b\Sigma Y = 0$

and dividing by N we have $\Sigma X/N - Na/N - b\Sigma Y/N = 0$

$$\text{or } M_X - a - bM_Y = 0$$

$$\begin{array}{l} \text{Substituting } X = x + M_X \quad \text{in } X = a + b Y \\ Y = y + M_Y \end{array}$$

$$\text{we get } x + M_X = a + by + bM_Y$$

Rearranging terms

$$\begin{aligned} x - by &= -M_X + a + bM_Y \\ &= -(M_X - a - bM_Y) \end{aligned}$$

where the quantity in brackets was previously shown to be equal to 0

therefore $x - b y = 0$, and $x = b y$

Similarly it can be shown that $y = b'x$

TECHNICAL APPENDIX B₂

To illustrate that any set of measurements, in terms of standard scores, has a mean of zero and a standard deviation of 1, consider the definitions of the mean and standard deviation:

$$M_X = \Sigma X / N$$

$$\sigma_X = \sqrt{\Sigma x^2 / N}$$

Therefore, the mean and standard deviation of standard scores are:

$$M_{z_X} = \Sigma z_X / N$$

$$= (\Sigma x / \sigma_X) / N$$

$$= \Sigma (x / N \sigma_X)$$

$$= \Sigma (X - M_X) / N \sigma_X$$

$$= (\Sigma X - N M_X) / N \sigma_X$$

Substituting $M_X = \Sigma X / N$,

$$M_{z_X} = [\Sigma X - (N/N) \Sigma X] / N \sigma_X$$

$$= 0 / N \sigma_X$$

$$= 0.$$

$$\sigma_{z_X} = \sqrt{\Sigma z_X^2 / N}$$

$$= \sqrt{\Sigma (x / \sigma_X)^2 / N}$$

$$= \sqrt{\Sigma x^2 / N \sigma_X^2}$$

$$= \sqrt{\sigma_X^2 / \sigma_X^2}$$

$$= 1$$

TECHNICAL APPENDIX B₃

The regression lines, in terms of deviation scores, have been previously shown to be

$$x = b y$$

$$y = b'x$$

Standard scores are defined as

$$z_X = x/\sigma_X$$

$$z_Y = y/\sigma_Y$$

so that

$$z_X \sigma_X = x$$

$$z_Y \sigma_Y = y$$

Substituting these values above, and rearranging terms

$$z_X \sigma_X = b z_Y \sigma_Y$$

$$z_X = (\sigma_Y/\sigma_X) b z_Y$$

Let $\sigma_Y/\sigma_X \quad b = B$

Then $z_X = B z_Y$

Similarly $z_Y = B' z_X$

which are the equations for the regression lines in terms of standard scores.

TECHNICAL APPENDIX B 4

The regression lines, in terms of standard scores, are:

$$z_X = Bz_Y$$

$$z_Y = B'z_X$$

where $B = B' = r_{XY}$. This may be demonstrated by considering the use of Bz_Y to predict z_X . This predicted value will have an error (e) where $e = z_X - Bz_Y$, and the sums of squares for all such errors will be

$$\sum_{1}^N e^2 = \sum (z_X - Bz_Y)^2$$

To make this sum a minimum, we must set the partial derivative with respect to B equal to zero and solve for B.

$$\begin{aligned} \frac{\partial \sum e^2}{\partial B} &= \frac{\partial \sum (z_X - Bz_Y)^2}{\partial B} \\ &= 2 \sum (z_X - Bz_Y) (-z_Y) \\ &= -2 \sum z_X z_Y + 2B \sum z_Y^2 \end{aligned}$$

But, since $\sigma^2_{z_Y} = \sum z_Y^2 / N = 1$,

then $\sum z_Y^2 = N$, and $\partial \sum e^2 / \partial B = -2 \sum z_X z_Y + 2BN$

Equating this value to 0, we get

$$-2 \sum z_X z_Y + 2BN = 0$$

$$2BN = 2 \sum z_X z_Y$$

$$B = \sum z_X z_Y / N$$

Similarly, using $B'z_X$ to predict z_Y , we get

$$\Sigma e^2 = \Sigma (z_Y - B'z_X)^2, \text{ and, finally}$$

$$\partial \Sigma e^2 / \partial B' = -2 \Sigma z_X z_Y + 2 B' N$$

$$\text{and } 2 B' N = 2 \Sigma z_X z_Y$$

$$B' = \Sigma z_X z_Y / N$$

And thus, since the correlation coefficient (r) is defined as

$$r_{XY} = \Sigma z_X z_Y / N, \text{ we have shown that } B = B' = r_{XY}.$$

TECHNICAL APPENDIX B 5

The average of the errors of prediction squared is equal to unity minus the square of the correlation coefficient. This follows from the definition of the error of prediction (e):

$$e = z_Y - r_{XY} z_X, \text{ or, } e = z_X - r_{XY} z_Y$$

Squaring and adding all such errors for the entire population (n), we get

$$\begin{aligned} \Sigma e^2 &= \Sigma (z_Y - r_{XY} z_X)^2 \quad \text{or} \quad \Sigma (z_X - r_{XY} z_Y)^2 \\ &= \Sigma z_Y^2 - 2 r_{XY} \Sigma z_X z_Y + r_{XY}^2 \Sigma z_X^2 \quad \text{or} \\ &\quad \Sigma z_X^2 - 2 r_{XY} \Sigma z_X z_Y + r_{XY}^2 \Sigma z_Y^2 \end{aligned}$$

but, recalling that

$$\Sigma z_Y^2 = \Sigma z_X^2 = N, \text{ and } \Sigma z_X z_Y = N r_{XY}$$

we have

$$\Sigma e^2 = N - 2N r_{XY}^2 + N r_{XY}^2, \text{ or, } N - 2N r_{XY}^2 + N r_{XY}^2$$

Dividing by N and combining terms, we get

$$\Sigma e^2 / N = 1 - r_{XY}^2 \text{ in either case.}$$

APPENDIX C

GLOSSARY OF ALGEBRAIC SYMBOLS

$A, B, \dots R$	Common reference abilities, or axes, or vectors
a a'	Intercepts of regression lines when stated in raw score form
B B'	Slope of regression lines when stated in standard scores, also equal to correlation coefficient (r_{XY})
b b'	Slope of regression lines when stated in deviation or raw score form
$E_1, E_2 \dots E_n$	Error reference axes (representing the unreliable portions of the individual measurements)
e	Error of prediction
h_i^2	Communality coefficient of a test, equals the sum of the squares of all factor loadings on a test
L_{iJ}	Factor loading or correlation of i th test with the J th factor or reference axis
M_X M_Y	Arithmetic mean, or average
r_{XY}	Pearson product moment correlation coefficient
$S_1, S_2, \dots S_n$	Specific reference axes representing reliable but not common portions of the individual measurements

σ_X	Standard deviation, square root of the va-
σ_Y	riance (σ_X^2)
u_i^2	Uniqueness of a test, equals $1 - h_i^2$
X	Original score
Y	
x	Deviation score, deviate value from mean
y	
z_i (where $i = X,$ Y, or 1, 2, ... n)	Standard score of an individual on a test or predictor
z_J (where $J = A,$ B, ... R)	Standard score of an individual on a ref- erence ability or factor

A P P E N D I X D

ACTIVITIES PARTICIPATING IN THE PROJECT

Conducted at the Medical Research Laboratory, this complex study involved participation by several contributing activities:

1. The Worcester Foundation for Experimental Biology where the ketosteroid and androgen assays were made.
2. The Grant Study, Harvard University, where the somatotyping and anthropometric measures were evaluated.
3. The Department of Psychology, Yale University, where a second scoring of the Rorschach data was undertaken.
4. The International Business Machine Co. of New Haven where large quantities of raw data were processed under the supervision of Medical Research Laboratory.
5. The Personnel Research Section of the Adjutant General's Office which maintained cognizance of the statistical procedures employed and provided assistance and advice.
6. The Watson Scientific Computing Laboratories, Columbia University, which offered its facilities for highly complex statistical work beyond the scope of ordinary IBM installations.

Within the Medical Research Laboratory itself, indoctrination of subjects and coordination of the test administration was handled by Captain C. W. Shilling (MC) USN, at that time officer-in-charge of the laboratory.

The physical fitness tests and physical examinations were administered by Commander J. J. Blanch (MC) USN

and Lt. Comdr. J. G. Bateman (MC) USN.

The somatotyping photographs were taken by Lt. Comdr. E. B. Cook (MSC) USN who also functioned as liaison officer with the various contributing activities.

The psychiatric interviews and psychological tests were conducted by Lt. A. H. Kaplan (MC) USNR, the laboratory psychiatrist, with the assistance of Lt. I. A. Everley (HC) USN.

Blood and urine counts were done by HMC L. C. Biggs and HMC A. S. Potopinski, and maintenance of the test schedule and control of the living conditions of the subjects were handled by HMC F. G. St. Louis and HMC C. M. Kennedy.

Lt. Comdr. E. B. Cook (MSC) USN was responsible for organization of the data and statistical analysis of results. Under his supervision, the work has been handled mainly by Mrs. Marian B. Elliott, Miss Inez A. Cox, Mrs. Nan L. Cook, Miss Doris Newman, HMC C. W. Tanner, HMC R. E. L'Etoile and HM3 H. D. Woolsey of the statistical section.

Guidance in final interpretation of results was furnished by the Medical Research Laboratory consulting statistician, Dr. R. J. Wherry of the Personnel Research Section, Adjutant General's Office, Department of the Army, and of the Department of Psychology, Ohio State University.

A complete list of Medical Research participants in the study is as follows:

E.S. = Experimental Staff

S.S. = Statistical Staff

Bateman, J. G., Lt. Comdr. (MC) USN	E.S.
Biggs, L. M., HMC, USN	E.S.
Blanch, J. J., Comdr. (MC) USN	(in charge) E.S.
Cook, E. B., Lt. Comdr. (MSC) USN	E.S.
	(in charge) S.S.

Cook, N. L., Civil Service	S.S.
Cox, I. A., Civil Service	S.S.
DeLashmutt, J. H., HM1, USN	S.S.
Elliott, M. B., Civil Service	S.S.
Everley, I. A., Lt. (HC) USN	E.S.
Kaplan, A. H., Lt. (MC) USNR	E.S.
Kennedy, C. M., HMC, USN	E.S.
L'Etoile, R. E., HMC, USN	S.S.
Mann, M. B., HMC, USN	S.S.
Morais, A. A., QMC, USN	E.S.
Newman, D., Civil Service	S.S.
Potopinski, A. S., HMC, USN	E.S.
Ryan, A., Civil Service	S.S.
Shilling, C. W., Captain (MC) USN (coordinator)	E.S.
St. Louis, F. G., HMC, USN	E.S.
Stover, A. H., HMC, USN	S.S.
Tanner, C. W., HMC, USN	S.S.
Westphal, D., HM3, USN	S.S.
Woolsey, H. D., HM3, USN	S.S.

A P P E N D I X E

ACKNOWLEDGMENTS

The present opportunity is taken to express acknowledgment to the many individuals and activities who have contributed to the success of these investigations:

To Captain C. W. Shilling (MC) USN, Officer-in-Charge of Medical Research Laboratory at the time of the initiation of the study. It was he who conceived the project as a whole and his continued interest and backing rendered the investigations possible.

To Captain T. L. Willmon (MC) USN, present Officer-in-Charge of Medical Research Laboratory. Assuming command in the midst of the analysis of this complex mass of data, his sympathetic appreciation of the time-consuming statistical processes involved and his faith in the ultimate worthwhile results of the study have withstood a year of concentrated computational work without interruption for interim reports.

To the many who contributed advice, equipment and background material, particularly Dr. Hudson Hoagland and Dr. Gregory Pincus, co-directors of the Worcester Foundation for Experimental Biology; to Dr. Carl Seltzer and Mr. David DeHarport of the Grant Study, Harvard University; to Dr. Edwin Henry, Dr. Donald Baier, Dr. Erwin Taylor and Mr. Arthur Fifer of the Personnel Research Section, Adjutant General's Office; to Dr. W. J. Eckhardt and Dr. H. R. J. Grosch of the Watson Scientific Computing Laboratories, Columbia University; to Dr. Jack Dunlap, then of the Psychological Corporation; to Captain C. O. Triebel, USN, and Dr. Leonard Mead, then of the Navy's Special Devices Center, Sands Point, L. I.; to Mrs. Edith Lisansky and Dr. Seymour Sarason of the Department of Psychology, Yale University; to Dr. Ross A. McFarland of the Harvard University School of Public Health; to Captain E. G. Hakan'sson (MC) USN of

the Naval Medical Research Institute, Bethesda, Md.; to Dr. Elliott Rodnick of the Worcester State Hospital; to Dr. Harrison Clarke of Springfield College; to the many Submarine Base departments who cooperated in the provision of equipment, subjects and special facilities, particularly Captain F. B. WarderUSN, then officer-in-charge of the Submarine School, and Captain Clark W. Grazier (MC) USN then head of the Base Dispensary; to the Medical Research group who functioned as a smoothly working unit throughout the experimental period, and to the statistical section who have labored in the analysis of data.